



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/608,587	06/27/2003	James M. Sweet	D/A2555	8422
25453	7590	04/11/2007	EXAMINER	
PATENT DOCUMENTATION CENTER			HILLERY, NATHAN	
XEROX CORPORATION			ART UNIT	PAPER NUMBER
100 CLINTON AVE., SOUTH, XEROX SQUARE, 20TH FLOOR			2176	
ROCHESTER, NY 14644				
SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE		
3 MONTHS	04/11/2007	PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

<b>Office Action Summary</b>	Application No.	Applicant(s)
	10/608,587	SWEET ET AL.
	Examiner Nathan Hillary	Art Unit 2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on 05 February 2007.  
 2a) This action is FINAL.      2b) This action is non-final.  
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1-37 is/are pending in the application.  
 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.  
 5) Claim(s) \_\_\_\_\_ is/are allowed.  
 6) Claim(s) 1-37 is/are rejected.  
 7) Claim(s) \_\_\_\_\_ is/are objected to.  
 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.  
 10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.  
     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).  
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
 a) All    b) Some \* c) None of:  
     1. Certified copies of the priority documents have been received.  
     2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
     3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |  |   |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                   | Paper No(s)/Mail Date. _____  |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
|  | 6) <input type="checkbox"/> Other: _____                                    |

## DETAILED ACTION

1. This action is responsive to communications: Amendment filed on 02/05/07.
2. Claims 1 – 37 are pending in the case. Claims 1, 16 and 27 are independent.

### ***Claim Rejections - 35 USC § 101***

3. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or, any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4. Claims 1 – 37 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Claims 1 – 37 have no practical application of a judicial exception as claimed because there is no physical transformation and no production of a concrete, useful and tangible result.
  - a. The claimed invention remains in the abstract and nothing is made available to the user; thus it does not produce a tangible result.
  - b. The claims appear to be in the preliminary stages and fall short of the disclosed practical utility. In other words, the claims fail to fulfill and/or reflect the specific, substantial, and credible utility sought by the disclosed invention, and thus do not produce a useful result.
5. Consequently, the claims are nonstatutory. The claims simply recite methodologies for assembling and grouping data without producing a concrete, useful, and tangible result.
6. Further, to expedite a complete examination of the instant application the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to make them statutory.

***Claim Rejections - 35 USC § 112***

7. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

8. Claims 1 – 37 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

9. The original specification as filed provides no support for a document representation stored in memory (penultimate lines of claims 1, 16 and 27).

10. Claims 2 – 15, 17 – 26, and 28 – 37, the claims are rejected for fully incorporating all of the deficiencies of the base claim(s) from which they depend.

***Claim Rejections - 35 USC § 103***

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

12. Claims 1 – 6, 10 – 13, 16 – 20, 25 – 31, 36, and 37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (6112203) and further in view of Earl (5924104).

13. **Regarding independent claim 27**, Bharat et al. teach that we locate pages that point to at least one of the pages in the start set 201. We call this set of pages the back set 202 (Column 4, line 61 – Column 5, line 20), which meets the limitation of **performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page.**

Bharat et al. teach that if a link points to a page that is represented by a node in the graph, and both pages are on different servers, then a corresponding edge 213 is added to the graph 211. Nodes representing pages on the same server are not linked. This prevents a single Web site with many self-referencing pages to unduly influence the outcome (Column 4, line 61 – Column 5, line 20), which meets the limitation of **identifying possible progression links; identifying possible table of content links; and examining the possible progression links and the possible table of content links for common characteristics**, since the specification states that “the possible progression links 230 and possible table of content links 240 are passed to module 250 for a final examination to weed out links which have properties that are not characteristic of typical intra-document links, e.g. they point to a different web server” (p 7, lines 26 – 30). It should be noted that pages on the same server are nodes and are thus still apart of the resulting graph.

Furthermore, it should be noted that the self-referencing pages of Bharat et al. are equivalent to intra-document links and that those intra-document links can be “possible” progression and/or table of contents links, since the Office has interpreted the word “possible” as “could be” and within the broadest, reasonable interpretation in light

Art Unit: 2176

of the specification, which states that a link analysis phase consists of the identification for a given hypertext page of the most likely desirable intra-document links. Those intra-document links fall into two categories: progression links and table of contents links (p 5, second paragraph). Thus, any intra-document link – a link that points to a different web server – could be a possible progression or table of contents link.

Bharat et al. teach that a larger n-graph 211 can be constructed by repeating this process for the back and forward sets 202-203 to add more indirectly linked pages (Column 4, line 61 – Column 5, line 20), which meets the limitation of **performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective set of identified candidate document pages is assembled.**

Bharat et al. do not explicitly teach performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents; examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of typical intra-document links, to provide a resultant set of identified candidate document pages; and grouping the content found in the resultant set of candidate document pages into a document representation stored in memory for subsequent viewing or printing by a user of the given hyperdocument.

Earl teaches that the link display manager 300 includes a display system for defining predetermined screen element properties providing visual cues for distinguishing the identified links 202 and 204. When a user provides an input link selection to select a new document, the document parser 304 parses the selected new document to identify intradocument links 202 and interdocument links 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages.**

Earl teaches that the display system 306 processes the identified intradocument links 202 and interdocument links 204 for displaying distinctively the intradocument links 202 and interdocument links 204 with predetermined visual cues to differentiate the links 202, 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **grouping the content found in the resultant set of candidate document pages into a document representation stored in memory for subsequent viewing or printing by a user of the given hyperdocument.**

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the invention of Bharat et al. with that of Earl because such a combination would provide the users of Bharat et al. with *an improved method and apparatus for displaying links on a user display interface in a computer system* (Column 1, lines 39 – 41).

14. **Regarding dependent claims 28 – 31,** Bharat et al. teach that the nodes in the start set are first scored according to their connectivity, and the number of terms of the query that appear as unique sub-strings in the URL of the represented documents. The score is a weighted sum of the number of directed edges to and from a node and the number of unique sub-strings of the URL that match a query term (Column 3, lines 10 – 15), which meet the limitation of **the page-level link analysis includes examination of contextual clues, the contextual clue is a particular class of content item associated with the hyperlink, the class of content item is a class of text, the class of text is a directional word or phrase.**

15. **Regarding dependent claims 36 and 37,** Bharat et al. teach that we assign a similarity weight to each node 213 of the sub-graph 255. Various document similarity measuring techniques have been developed in Information Retrieval to determine the goodness of fit between a "target" document and a collection of documents. These techniques typically measure a similarity score based on word frequencies in the collection and a target document (Column 6, lines 51 – 57), which meet the limitation of **the page-level analysis includes determining the similarity of the hyperlink destination to that of other hyperlinks within the page, and the page-level analysis includes determining the similarity of the hyperlink destination to the location of the current page.**

16. **Regarding claims 1 – 6 and 10 – 13,** the claims incorporate substantially similar subject matter as claims 27 – 31, 36 and 37 and are rejected along the same rationale.

17. **Regarding claims 16 – 20, 25 and 26,** the claims incorporate substantially similar subject matter as claims 27 – 31, 36 and 37 and are rejected along the same rationale.

18. Claims 7 – 9, 14, 15, 21 – 24, and 32 – 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (5924104) as applied to claims 1 – 6, 10 – 13, 16 – 20, 25 – 31, 36, and 37 above, and further in view of Prince (US 6877002 B2).

19. **Regarding dependent claims 32 – 35,** neither Bharat et al. nor Earl explicitly teach **the class of content item is a class of image, the class of image is an image containing a directional symbol, a textual clue is obtained for the image, the identifying of table of content links includes the presence of at least one other hyperlink nearby with the page description.**

However, Prince teaches that the parsed results (from step 42 in FIG. 4) relating to the media are passed to extraction agent 68 via an extraction queue 67. The extraction queue 67 comprises URLs to be analyzed with respect to associated media metadata. The extraction queue 67 may comprise metadata queue entries such as media URLs, Web page URLs, Web page titles, Web page keywords, Web page descriptions, media title, media author, and media genre. Each queue entry added to

the extraction queue is assigned a processing time and a priority (Column 7, lines 23 – 37), meets the limitation of **the class of content item is a class of image, the class of image is an image containing a directional symbol, a textual clue is obtained for the image, the identifying of table of content links includes the presence of at least one other hyperlink nearby with the page description.**

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the combined invention of Bharat et al. and Earl with that of Prince because such a combination would allow the users of Bharat et al. and Earl the benefit of *A method for querying metadata associated with media on a computer network includes separating the metadata into keywords* (Column 2, lines 37 – 39).

20. **Regarding claims 7 – 9, 14 and 21 – 24,** the claims incorporate substantially similar subject matter as claims 32 – 35 and are rejected along the same rationale.

21. **Regarding claim 15,** Bharat et al. teach that we assign a similarity weight to each node 213 of the sub-graph 255. Various document similarity measuring techniques have been developed in Information Retrieval to determine the goodness of fit between a "target" document and a collection of documents. These techniques typically measure a similarity score based on word frequencies in the collection and a target document (Column 6, lines 51 – 57), which meet the limitation of **the page-level analysis includes determining the similarity of the hyperlink destination to that of other hyperlinks within the page.**

***Response to Arguments***

22. Applicant's arguments filed 2/5/07 have been fully considered but they are not persuasive.
23. Applicant argues that claims 1 – 37 are statutory under 35 USC 101 because a document stored in memory is not a judicial exception, and it has a transformation (p 9).

The Office disagrees.

First, a document stored in memory, whether inside someone's brain or on a computer, constitutes a judicial exception; specifically, it is considered an abstract idea. Secondly, the transformation to which applicant eludes, web page data into a document representation stored in memory, simply constitutes a data transformation not a physical transformation.

Further, applicant subtly requests suggestions on how to overcome the 101 rejection. To this end, it is suggested that applicant amend each independent claim to make them statutory by producing a tangible and useful result. For example, positively reciting that the document is printed, stored or displayed NOT insinuating that it might happen some time in the future and NOT simply reciting an intended use, future or otherwise, for the document. It should be noted that the suggestions mentioned, even if made, may not meet the standards under 35 USC 112, first paragraph.

24. Applicant argues that Earl does not teach examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of

**identified candidate document pages** because what Earl defines as intra-document is what Applicant would call intra-page and thus what Earl calls inter-document is really inter-page (p 12, second full paragraph).

The Office disagrees.

First, nowhere in the original Specification does Applicant define or even discuss "intra-page" or "inter-page". Second, Applicant proffers a "definition" of intra-document taken from the Specification (p 11, last paragraph). "The first step for an automated system for the identification of multi-page documents is to identify links within a given web page that may link to other pages within the same document. Such links are referred to as intra-document links". This "definition" is clearly NOT limiting in any way.

Second, the claim recites **links, which have properties that are not characteristic of intra-document links**. Consequently, the Office is forced to rely upon the knowledge of one of ordinary skill in the art in order to interpret the broad limitation recited in light of the lack of a definitive definition in the Specification. Thus, the Office maintains that Earl clearly and explicitly teaches intra-document and inter-document links that meet the claimed intra-document link in spite of Applicant's attempts to split hairs without sufficient evidence to support it.

25. Applicant argues that Earl further does not teach **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages** because Earl discriminates between two types

of links but keeps all those links while applicant discards or weeds out those links (p 12, last paragraph).

The Office disagrees.

First, Applicant is correct that Encarta defines "weed out" as "to separate out something undesirable. However, the Office maintains that Earl does "weed out" the links within the broadest, reasonable interpretation in light of the specification. The term "weed out" is not defined in the specification. Although the applicant attempts to explain what the term "weed out" should mean as it pertains to gardening, the Office is forced to rely on the knowledge of one of ordinary skill in the art NOT of gardening but of computer technology.

Thus, Earl, by applicant's own admission, teaches discriminating visually between intra-document and inter-document links, which meet the definition of separating out, or weeding out, the links visually on screen. The requirement to have to discard the links is too limiting in view of what is actually claimed.

### ***Conclusion***

26. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan Hillery whose telephone number is (571) 272-4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on (571) 272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

  
Doug Hutton  
Primary Examiner  
Technology Center 2100

NH